

# The Place of Epidemiology in Environmental Decisions: Needed Support for the Development of Risk Assessment Policy

Raymond Richard Neutra<sup>1\*</sup> and Dimitrios Trichopoulos<sup>2</sup>

<sup>1</sup>California Department of Health Services, 5900 Hollis St., Suite 8, Emeryville, CA 94704; <sup>2</sup>Department of Epidemiology, Harvard School of Public Health, 667 Huntington Avenue, Boston, MA 02115

Some of the most challenging problems in the use of epidemiology for regulatory policy concern summarizing epidemiological and other kinds of information to create a weight of evidence. Another frequent issue is whether to embark on epidemiological study. There are also concerns that negative results never see the light of day. These and other meta-issues are worthy of funded evaluation by expert work groups. — *Environ Health Perspect* 101(Suppl 4): 67–69 (1993).

Key Words: Epidemiology, risk assessment policy, research needs

## Risk Assessment Policy Needs Go Beyond Research Needs

So far we have discussed areas of exposure analysis, study design, and data analysis in which methodological improvements are needed and worthy of research support. For the most part, this support will go to individual researchers to help finance the time spent developing and testing new methods.

In this paper we discuss scientific activities, such as the systematic and consistent summarization of a body of evidence, the decision to initiate more study, and the conclusion that enough is enough.

These are issues of risk assessment policy and research strategy, and while not as value laden as risk management (which we do not address in this document), they are by nature a matter of scientific consensus. For this reason, support for the formation of a risk assessment policy often involves development of draft documents by governmental and nongovernmental scientists and the systematic review and development of procedure by groups of researchers. We begin by raising some of the issues of a risk assessment policy and discuss in the last section how this policy or related research needs could be supported. To place these considerations in context, we must remember that most risk assessment must proceed with only animal data. Therefore, the issues raised here, though important, apply to a small proportion of regulatory decisions.

## Putting It All Together: Methodological Issues in Weighing a Body of Evidence

Repeated strong epidemiological findings can implicate a remediable environmental exposure even without supporting animal toxicological evidence or identifying a responsible agent. An example would be the well-known strong association between carcinoma of the nasal sinus and cabinet making (1). Industrial hygienic precautions can be instituted even before a better understanding of the responsible mechanism is clarified. Here, epidemiology alone suffices to drive regulation. Strong consistent results require no ingenuity to summarize.

More frequently, the human epidemiological results are not distinguishable from the null, or the dose-response slope is so low that bias or confounding could plausibly account for the observed association. Alternatively, the results may implicate something in the general environment that cannot be avoided by some easy measure that could, like a cabinetmakers dust mask, be applied without an understanding of the responsible agent. In this case, usually other disciplines are needed to pinpoint the offending agent and its means of control. For all of these reasons, most environmental policy is set after considering the integrated information from clinical medicine, basic physical and biological science, animal toxicology, exposure analysis, and a body of epidemiological evidence. When all of this evidence points toward the same conclusion, policy decisions are simplified. Often, however, the evidence is conflicting. Some studies are called positive, and others are said to be negative, that is, indistinguishable from the null.

The terms positive and negative suggest a solidity that is misleading. One school of

thought reserves these terms for statistically significant associations and tends to view any association that does not achieve the preset  $p$  value to be as good as no association at all. Another school suggests that information from all studies should be pooled and the decision to believe the results should be based on Bayesian approaches that consider prior plausibility and the cost of a false positive or a false negative result. This issue becomes particularly difficult when, as is the case in the current debate about low-frequency electromagnetic fields (2), the body of positive epidemiological evidence has very weak biological plausibility. Intuitively it is clear that a higher relative risk or a greater number of confirmatory studies are necessary for such a situation than would be for an agent that is similar to a previously studied agent whose mechanism of action is well understood. The acceptance of erionite as a carcinogenic mineral fiber comes to mind. The documentation of two villages with a rate ratio of 9000 and one animal study demonstrating carcinogenicity was sufficient for the International Agency for Research on Cancer (IARC) (3) to list erionite as a carcinogen. The biological plausibility weighed heavily here. Methodological research into how one uses epidemiological and other types of information to update prior probability assessments practically and intelligibly for environmental decision makers is, therefore, of high priority.

Agents that act by nonthreshold mechanisms, as is the case with many carcinogens, could conceivably produce unacceptable numbers of disease in the population when there is widespread exposure. This could happen even when the exposures are so low that the increase in relative risk is undetectable.

This manuscript was prepared as part of the Environmental Epidemiology Planning Project of the Health Effects Institute, September 1990–September 1992.

\*Author to whom correspondence should be addressed.

Although epidemiological studies cannot rule out effects of societal concern in this situation, such negative or null studies can sometimes help assess whether humans have higher or lower sensitivity to an agent than expected on the basis of animal bioassays. This is a special topic within the problem of summarizing evidence.

Another problem that often faces those who are trying to weigh a body of evidence is that negative studies (not distinguishable from the null) are thought to be less likely to be published than positive studies. Censoring may occur at the level of the researcher or the journal. For example, a researcher may suggest a positive association between a disease outcome and a variable that was not originally included in the main hypothesis. The researcher probably would not report a surprising lack of association under the same circumstances. An author whose main hypothesis was not supported in a study may decide not to submit an article or may become discouraged after receiving routine editorial criticisms. All of this could skew the available evidence. This problem has been discussed previously (4). It is time to see if the phenomenon is substantial and to evaluate which, if any, potential remedies should be applied.

### Contexts in which Follow-up Environmental Epidemiologic Studies are Recommended

The usual motivation for academically based epidemiologic research is to pursue a credible hypothesis in a setting that promises a high likelihood of providing a persuasive answer because the amount of exposure, the size of the study population, and the ability to control bias and confounding are all favorable. Research priorities from funding agencies thus give as much attention to feasibility as to the potential importance of the project being funded. Hypotheses often derive from basic science considerations or from animal or other epidemiological evidence. Public agencies, on the other hand, are often directed to carry out studies whose answer would be of great policy interest even though the low biological credibility of the tested hypothesis or the conditions of study militate against the likelihood of a persuasive result.

Frequently, one or a few studies initiated in either way are not considered persuasive. When current epidemiological information is insufficient, one is faced with the problem of deciding if any additional epidemiological study is likely to be helpful (for example, where animal evidence suggests the possibility of an effect large enough to be of social concern but too small to be detected toxicologically or epidemiologically under usual

exposure scenarios). If epidemiology offers hope for demonstrating an effect, how strong must a collection of positive studies be to implicate an agent or estimate its potency in humans? How strong must a collection of negative studies be to give a clean bill of health to an agent that is thought to act by a mechanism that should display a threshold of effect? In regulatory toxicology, no single study is ever considered definitive. Instead, a specified number of statistically significant studies in several species and laboratory settings is routinely required. This policy is pursued regardless of considerations such as strength of effect, prior plausibility, or the social costs of false positive or negative results. While this procedure would not be advocated for epidemiology, it does remind us that no single study, especially a screening study, is likely to be definitive and that initiating an investigation for scientific or public policy reasons usually commits one to a sequence of studies until a body of evidence has accumulated. The principles that guide the initiation and termination of this commitment need elucidation.

When concerned segments of the public demand a study in a particular setting, for instance the Love Canal or large areas of Los Angeles where there is aerial application of malathion for the Mediterranean fruit fly, they often have a legitimate desire to participate in the design, conduct, and interpretation of the results. They often have unique "shoe leather" experience as to routes of exposure, hypotheses about effects that they wish to be addressed, and susceptibility to and concerns about potential conflicts of interest in the analysis process. Often the results (or media interpretation of the results) from one such location can have profound impact on national policy.

Several methodological research questions arise from the involvement of members of the public. How well do various techniques of involvement work, practically, for citizen satisfaction, and with regard to avoiding bias in study results? How does the presentation of epidemiological results and the social setting in which that presentation takes place influence the understanding and acceptance of results? How can one safely assume that a local community correctly registers and remembers study results carried out on its behalf, given that its perception of the results influences local and national policy?

The availability of morbidity and mortality data bases offers both opportunities and dilemmas. Academic epidemiology has traditionally turned to available data bases to look for unsuspected variations with regard to person, place, and time as a first step to unearthing the

activity of some causal agent. This approach was often useful in the early days of infectious disease epidemiology; and recently there have been some successes in China, where sharp regional variations in chronic disease rates (5) have led to the discovery of causal agents of indoor air pollution [female lung cancer in southern China (6)] or deficiencies or excesses in trace elements [selenium (7)].

Despite these isolated instances and the usefulness of these data bases in occupational and life-style epidemiology, it is striking how difficult it is to find examples in modern developed countries in which the routine surveillance of morbidity and mortality for temporal or spatial variations has led to the discovery of new causal factors in the physical environment. These sources of information have been more useful for following the course and control of disease of known origin and for testing specific hypotheses that arose from other considerations. Despite this unpromising record, there are politicians and scientists who have high hopes for screening studies based on the analysis of routine data. This can be fruitful even when there is no biologically based hypothesis, but the opportunity for chance associations from multiple comparisons is great. In a more hypothesis-testing mode, one can explore existing data to ask whether a single locality (or all localities) containing some environmental hazard has (or have) higher rates of a particular disease. The opportunities for chance associations after multiple comparisons also are great. The availability of such data also facilitates the generic clustering study in which one asks if a disease clusters more than chance or demographics would suggest. The answer allows the researcher to assess the hypothesis that a disease is spread from person to person or from a point source. The traditional academic hypothesis-generating strategies thus have potential for initiating a wild-goose chase in environmental epidemiology. There is a need for a clear rationale for initiating such endeavors while considering the subsequent commitment it entails.

### Methodological Research Questions

We have raised some research and policy issues that arise when summarizing epidemiological evidence or deciding to initiate new studies in service to the regulatory process. We will briefly discuss them below.

How can funding agencies advance the state of knowledge and the quality of practices in these several areas? As mentioned before, science policy requires a consensus process and, therefore, requires the support of researchers working as a group. Efficiency dictates that some researchers be supported to prepare the

groundwork or oversee and summarize the consensus process while others are supported to participate in it. Care should be given to inviting participants with a range of disciplines and backgrounds so that persons with various kinds of practical experience are balanced by individuals who are not grounded in the more traditional ways of doing or thinking about things. A variety of mechanisms could be appropriate for the different areas dealt with below. These include supporting outside scientists to work with government scientists to develop draft guidelines or work out discussions of the rationale for approaching a particular problem. They include RFP support for individuals, the intensive work group sessions to resolve a problem procedure, or a workshop or conference to summarize a consensus. While support for such activities can be as costly as support for laboratory or clinical research, the yield can be useful for the regulatory and scientific processes.

### Epidemiological Evidence

How should a body of epidemiological evidence be summarized for hazard identification and dose-response purposes? How should biological background information be incorporated?

These questions concern more than statistical meta-analysis. There are some very interesting issues relating to the ability of the human mind to integrate large bodies of information and summarize them in consistent ways. Prior opinions often influence the interpretation and weight of evidence. This area may benefit from an analysis of expert behavior and artificial intelligence. The issue of positive and negative studies arises. It is inappropriate to pass each study through a test of statistical significance and contrast the number of positive and negative studies. What alternatives are there to the semantics of this terminology and the practices that arise from them? To the extent that the process becomes more explicit, hidden biases will be less important and the process will be less arbitrary and capricious for regulatory purposes.

### The Price of Sample Information

When should epidemiological study be supported rather than laboratory or clinical study, and how many studies of varying kind and size are needed to determine that a hazard exists at current doses or that a hazard is unlikely at existing doses?

We cannot hope for a cut-and-dried procedure to make these determinations, but the elements that should go into such decisions need broader and more fundamental discussion and understanding. We need to bring together decision analysts, epidemiologists, statisticians, regulatory lawyers, and toxicologists to study case histories and to propose theoretically sound and practical approaches to deciding when to start and how much is enough. A published discussion of the issues, similar to the National Research Council's (NRC) risk assessment/risk management report of 1983 (8), may reduce the frequency of false positive and negative epidemiological studies, inappropriate waiting for unneeded studies, and the misuse of an inadequate number of negative or positive studies to make decisions.

### Documenting and Remediating Negative Publication Bias

Despite thoughtful discussions of the allegation that negative studies are less likely to be published, the extent of the problem and reasons (if any) for it have not been fully documented. Once this is better understood, along with the likely uses of negative results, alternative approaches should be proposed and a market survey conducted to determine the acceptability of the alternatives. A journal of abstracts is one option, and a peer-reviewed electronic data base with abstracts keyed into electronically retrievable detailed documentation is another. The benefit of understanding and remediating this problem is that negative studies, and reviewers' concerns about them, would be available as societal decisions are being made. However, an obvious danger is that negative studies may be negative because they are poorly conceived or poorly executed.

### The Implications of Involving the Public

What evidence claims that public concern or involvement can improve or bias study results or influence response rates? Once a study is done, what is the prevalence of knowledge about its results, and what is the duration of that knowledge? What interventions influence this? Here lies the borderland between epidemiology and evaluation research, similar to research in antismoking or contraception campaigns, except that segments of the public may view effective efforts to disseminate epidemiological research information as propaganda for environmental inaction. The benefits of a better understanding would be improved public health practice and public decision making based on evidence rather than misconception. A caveat must be given here. Environmentally cautious actions are often warranted even in the face of negative epidemiological results that are well disseminated and understood by the public. It is nonetheless helpful to be clear, for future reference, as to whether epidemiological evidence influenced the decision.

### Why Epidemiology Must Be Skeptical

Theory can predict that a certain proportion of comparisons will be statistically significant or that a shared methodological flaw can produce a false association in multiple studies. But both the scientific community and the general public are more convinced by empirical demonstrations of these theoretical predictions. For example, in how many occupational studies is a variable like month of birth associated with disease? How many examples can we find in which several studies seemed to implicate an agent but subsequent studies failed to confirm the initial findings? Case studies like these can help journalists, the public, and decision makers to better understand the importance of a solid body of evidence for making good policy. ♣

### REFERENCES

1. IARC. Monographs on the evaluation of carcinogenic risks to humans, Supplement 7. Lyon: International Agency for Research on Cancer, 1987; 380-381.
2. Office of Technology Assessment. Biological effects of power frequency electric and magnetic fields. Stock Number 052-003-01152-2. U.S. Government Printing Office, 1989.
3. IARC. Monographs on the Evaluation of Carcinogenic Risks to Humans, Supplement 7. Lyon: International Agency for Research on Cancer, 1987; 203-204.
4. Begg CB, Berlin JA. Publication bias: a problem interpreting medical data. *J Royal Soc Series A* 151:419-463 (1988).
5. Li JY, Liu PQ, Li GY, Rong S, Cao D, eds. Atlas of cancer mortality in the People's Republic of China, ed 1. Shanghai, China: China Map Press, 1979.
6. He X, Wei C, Liu Z, Chapman R. An epidemiological study of lung cancer in Xuan Wei County. *Environ Health Perspect* 94:9-13 (1991).
7. Keshan Disease Research Group. Epidemiological studies on the etiological relationship of selenium and Keshan disease. *Chin Med J* 92:477-482.
8. National Research Council. Risk assessment in the federal government: managing the process. Washington, DC: National Academy Press; 196-203.